

Mervyn G. Marasinghe • Kenneth J. Koehler

Statistical Data Analysis Using SAS

Intermediate Statistical Methods

Second Edition

 Springer

Contents

1	Introduction to the SAS Language	1
1.1	Introduction	1
1.2	Basic Language: A Summary of Rules and Syntax	8
1.3	Creating SAS Data Sets	13
1.4	The INPUT Statement	16
1.5	SAS Data Step Programming Statements and Their Uses	21
1.6	Data Step Processing	31
1.7	More on INPUT Statement	39
1.7.1	Use of Pointer Controls	39
1.7.2	The trailing @ Line-Hold Specifier	41
1.7.3	The trailing @@ Line-Hold Specifier	43
1.7.4	Use of RETAIN Statement	44
1.7.5	The Use of Line Pointer Controls	46
1.8	Using SAS Procedures	48
1.9	Exercises	59
2	More on SAS Programming and Some Applications	69
2.1	More on the DATA and PROC Steps	69
2.1.1	Reading Data from Files	70
2.1.2	Combining SAS Data Sets	72
2.1.3	Saving and Retrieving Permanent SAS Data Sets	78
2.1.4	User-Defined Informats and Formats	82
2.1.5	Creating SAS Data Sets in Procedure Steps	89
2.2	SAS Procedures for Descriptive Statistics	94
2.2.1	The UNIVARIATE Procedure	98
2.2.2	The FREQ Procedure	105
2.3	Some Useful Base SAS Procedures	122
2.3.1	The TABULATE Procedure	122
2.3.2	The REPORT Procedure	129
2.4	Exercises	139

3	Introduction to SAS Graphics	147
3.1	Introduction	147
3.2	Template-Based Graphics (SAS/ODS Graphics)	151
3.3	SAS Statistical Graphics Procedures	155
3.3.1	The SGPLOT Procedure	156
3.3.2	The SGPANEL Procedure	173
3.3.3	The SGSCATTER Procedure	182
3.4	ODS Graphics from Other SAS Procedures	186
3.5	Exercises	193
4	Statistical Analysis of Regression Models	199
4.1	An Introduction to Simple Linear Regression	199
4.1.1	Simple Linear Regression Using PROC REG	201
4.1.2	Lack of Fit Test	207
4.1.3	Diagnostic Use of Case Statistics	208
4.1.4	Prediction of New y Values Using Regression	217
4.2	An Introduction to Multiple Regression Analysis	221
4.2.1	Multiple Regression Analysis Using PROC REG	225
4.2.2	Case Statistics and Residual Analysis	231
4.2.3	Residual Plots	236
4.2.4	Examining Relationships Among Regression Variables	243
4.3	Types of Sums of Squares Computed in PROC REG	248
4.3.1	Model Comparison Technique and Extra Sum of Squares	248
4.3.2	Types of Sums of Squares in SAS	250
4.4	Subset Selection Methods in Multiple Regression	254
4.4.1	Subset Selection Using PROC REG	261
4.4.2	Other Options Available in PROC REG for Model Selection	272
4.5	Model Selection Using PROC GLMSELECT: Validation and Cross-Validation	273
4.6	Exercises	282
5	Analysis of Variance Models	301
5.1	Introduction	301
5.1.1	Treatment Structure	304
5.1.2	Experimental Designs	305
5.1.3	Linear Models	306
5.2	One-Way Classification	308
5.2.1	Using PROC ANOVA to Analyze One-Way Classifications	317
5.2.2	Making Preplanned (or A Priori) Comparisons Using PROC GLM	325
5.2.3	Testing Orthogonal Polynomials Using Contrasts	331
5.3	One-Way Analysis of Covariance	337
5.3.1	Using PROC GLM to Perform One-Way Covariance Analysis	339

5.3.2	One-Way Covariance Analysis: Testing for Equal Slopes	347
5.4	A Two-Way Factorial in a Completely Randomized Design	355
5.4.1	Analysis of a Two-Way Factorial Using PROC GLM	358
5.4.2	Residual Analysis and Transformations	363
5.5	Two-Way Factorial: Analysis of Interaction	367
5.6	Two-Way Factorial: Unequal Sample Sizes	375
5.7	Two-Way Classification: Randomized Complete Block Design	386
5.7.1	Using PROC GLM to Analyze a RCBD	389
5.7.2	Using PROC GLM to Test for Nonadditivity	395
5.8	Exercises	398
6	Analysis of Variance: Random and Mixed Effects Models	419
6.1	Introduction	419
6.2	One-Way Random Effects Model	423
6.2.1	Using PROC GLM to Analyze One-Way Random Effects Models	426
6.2.2	Using PROC MIXED to Analyze One-Way Random Effects Models	430
6.3	Two-Way Crossed Random Effects Model	438
6.3.1	Using PROC GLM and PROC MIXED to Analyze Two-Way Crossed Random Effects Models	441
6.3.2	Randomized Complete Block Design: Blocking When Treatment Factors Are Random	448
6.4	Two-Way Nested Random Effects Model	449
6.4.1	Using PROC GLM to Analyze Two-Way Nested Random Effects Models	451
6.4.2	Using PROC MIXED to Analyze Two-Way Nested Random Effects Models	455
6.5	Two-Way Mixed Effects Model	457
6.5.1	Two-Way Mixed Effects Model: Randomized Complete Block Design	460
6.5.2	Two-Way Mixed Effects Model: Crossed Classification	471
6.5.3	Two-Way Mixed Effects Model: Nested Classification	482
6.6	Models with Random and Nested Effects for More Complex Experiments	494
6.6.1	Models for Nested Factorials	494
6.6.2	Models for Split-Plot Experiments	500
6.6.3	Analysis of Split-Plot Experiments Using PROC GLM	503
6.6.4	Analysis of Split-Plot Experiments Using PROC MIXED ..	509
6.7	Exercises	516

7	Beyond Regression and Analysis of Variance	529
7.1	Introduction	529
7.2	Nonlinear Models	529
7.2.1	Introduction	529
7.2.2	Growth Curve Models	531
7.2.3	Pharmacokinetic Application of a Nonlinear Model	537
7.2.4	A Model for Biochemical Reactions	543
7.3	Generalized Linear Models	549
7.3.1	Introduction	549
7.3.2	Logistic Regression	552
7.3.3	Poisson Regression	569
7.4	Generalized Linear Models with Overdispersion	574
7.4.1	Introduction	574
7.4.2	Binomial and Poisson Models with Overdispersion	576
7.4.3	Negative Binomial Models	582
7.5	Further Extensions of Generalized Linear Models	587
7.5.1	Introduction	587
7.5.2	Poisson Regression with Rates	588
7.5.3	Logistic Regression with Multiple Response Categories	598
7.6	Exercises	612
Appendix A	SAS Templates	621
A.1	Introduction	621
A.1.1	What Are Templates?	621
A.1.2	Where Are the SAS Default Templates Located?	624
A.1.3	More on Template Stores	627
A.2	Templates and Their Composition	628
A.2.1	Style Templates	630
A.2.2	Style Elements and Attributes	631
A.2.3	Tabular Templates	633
A.2.4	Simple Table Template Modification	635
A.2.5	Other Types of Templates	637
A.3	Customizing Graphs by Editing Graphical Templates	638
A.4	Creating Customized Graphs by Extracting Code from Standard Graphical Templates	641
Appendix B	Tables	645
References		671
Index		675